603799

√ gn

(1)

COPY 1 of 1 COPIES

BRANCHING PROCESSES


T.E. Harris

Douglas Aircraft Company, Inc.


P -14

Revised: 29 July 1948

Approved for OTS release

37 p $2.00 hc
$0.50 mf



The RAND Corporation

1700 MAIN ST. • SANTA MONICA • CALIFORNIA

603799

RAOP-14*
7/29/48

# BRANCHING PROCESSES

by

T. E. Harris

1. Summary. This paper is concerned with a simple mathematical model for a branching stochastic process. Using the language of family trees we may illustrate the process as follows. The probability that a man has exactly $r$ sons is $p_r$, $r = 0,1,2,\ldots$. Each of his sons (who together make up the first generation) has the same probabilities of having a given number of sons of his own; the second generation have again the same probabilities, and so on. Let $z_n$ be the number of individuals in the $n^{th}$ generation. We study the probability distribution of $z_n$. Some previous results are given in section 2; these include procedures for computing moments of $z_n$, and a criterion for when the family has probability 1 of dying out. In sections 3 and 4 the case is considered where the family has a nonzero chance of surviving indefinitely. In this case the random variables $z_n/Ez_n$ converge in probability to a random variable $w$ with cumulative distribution $G(u)$. It is shown that $G(u)$ is absolutely continuous for $u \neq 0$. Results of a Tauberian character are given for the behavior of $G(u)$ as $u \longrightarrow 0$ and $u \longrightarrow \infty$. In section 5 some examples are given where $G(u)$ can be found explicitly; $G(u)$ is computed numerically for the case $p_1 = 0.4$, $p_2 = 0.6$. In section 6 families with probability 1 of extinction are considered. A method is given for obtaining in certain cases an expansion for the moment-generating function of the number of generations before extinction occurs.

*Revised 7-29-48

In section 7 maximum likelihood estimates are obtained for the $p_r$ and for the expectation $Ez_1$; consistency in a certain sense is proved. In section 8 a brief discussion is given of the relation between two types of mathematical models for branching processes.

2. Introduction. By a branching stochastic process is meant a phenomenon of the following general type: each of an initial aggregate of objects can give rise to more objects of the same or different types; the objects produced can then produce more, and the system develops, subject to certain probability laws. Examples are the development of human or animal populations, propagation of genes, and nuclear chain reactions. The mathematical model dealt with in this paper may be thought of as representing the generation-by-generation growth of a family, the fundamental random variable being the number of individuals in the $n^{th}$ generation. Under certain conditions, however, this model may describe the size of a family at a sequence of points in time. This question will be touched on in section 8.

Definition 2.1. The random variables $z_n$, $n = 0,1,2,\ldots$, will be said to represent a simple discrete branching process provided: $z_o = 1$; $P(z_1=r) = p_r$, $r = 0,1,2,\ldots$, with $\sum_{r=0}^{\infty} p_r = 1$; the conditional distribution of $z_{n+1}$, given $z_n=r$, is that of the sum of $r$ independent random variables, each having the same distribution as $z_1$.

Assumptions. Throughout this paper we assume that $\sum_{r=o}^{\infty} r^2 p_r < \infty$, that at least two of the $p_r$ are positive, and that $p_o+p_1 < 1$.

Definitions 2.2. Let $x = Ez_1 = \sum rp_r$, $\sigma^2 = \mathrm{Var}(z_1) = \sum r^2p_r - x^2$. Let $f(s) = \sum_{r=o}^{\infty} p_r s^r$ be the generating function of $z_1$

(s denotes a complex variable.) Let $p_{nr} = P(z_n = r)$ and $f_n(s) = \sum_{r=0}^{\infty} p_{nr} s^r$; of course $p_{1r} = p_r$ and $f_0(s) = s$. The assumptions given above insure that the first and second derivatives $f'(s)$ and $f''(s)$ are continuous in the set consisting of the interior of the unit circle and the point $s = 1$; thus derivative notations such as $f''(1)$ are used even though $f(s)$ may not be analytic at $s = 1$. It will be seen shortly that a similar remark applies to the functions $f_n(s)$ and certain functions to be introduced later.

In the remainder of this section we shall summarize certain results; most of them are contained implicitly or explicitly in works by Fisher [1], Lotka [2], Steffensen [3], Ulam and Hawkins [4], Kolmogoroff [5], Kolmogoroff and Dmitriev [6], and Yaglom [7]; some of these references are not widely available.

From our definition, $P(z_{n+1} = k \mid z_n = j)$ is the coefficient of $s^k$ in $[f(s)]^j$. Hence $p_{n+1,k}$ is the coefficient of $s^k$ in $\sum_{j=0}^{\infty} p_{nj} [f(s)]^j$, whence

(2.1) $$f_{n+1}(s) = f_n[f(s)].$$

Letting $n = 1,2,\ldots$, successively, it follows that <u>the generating function of $z_n$ is the $n^{th}$ functional iterate of $f(s)$</u>. Hence

(2.2) $$f_{n+1}(s) = f[f_n(s)].$$

We note that $f_n'(1) = Ez_n$, $f_n''(1) + f_n'(1) - [f_n'(1)]^2 = \mathrm{Var}(z_n)$. Differentiation of (2.1) at $s = 1$ gives $f_{n+1}'(1) = x^{n+1}$; another differentiation

Revised 7-29-48

gives $f''_{n+1}(1) = f''(1)\left[f'_n(1)\right]^2 + f'(1)f''_n(1)$ while twofold differentiation of (2.2) gives $f''_{n+1}(1) = f''(1)f'_n(1) + \left[f'(1)\right]^2 f''_n(1)$; these two expressions for $f''_{n+1}(1)$ can be equated and solved for $f''_n(1)$, provided $x = f'(1) \neq 1$. Thus <u>the mean and variance of $z_n$ are given by</u>

$$Ez_n = (Ez_1)^n = x^n; \quad \mathrm{Var}(z_n) = \frac{\sigma^2 x^n(x^n-1)}{x^2-x},\ x \neq 1; \quad \mathrm{Var}(z_n) = n\sigma^2,\ x = 1.$$

Higher moments, if they exist, may be found by a similar process.

<u>Definition</u> 2.3. Denote by <u>a</u> the smallest non-negative real root of the equation $t = f(t)$. We see that $x \leq 1$ implies $a = 1$ while $x > 1$ implies $0 \leq a < 1$, the equality $a = 0$ holding if and only if $p_0 = 0$. In no case can the half-open interval $0 \leq t < 1$ contain more than one root. It is readily seen that

$$(2.3) \qquad \lim_{n\to\infty} p_{no} = \lim f_n(0) = a.$$

We thus have the well known result: <u>the number</u> a <u>is the probability of eventual extinction of the family</u>. The relation between <u>a</u> and <u>x</u> shows that <u>the probability of extinction is</u> 1 <u>if and only if</u> $x \leq 1$.

It is also clear that $0 \leq t < 1$ implies $\lim f_n(t) = a$; this, together with (2.3), shows that

$$(2.4) \qquad \lim_{n\to\infty} p_{nr} = 0, \quad r = 1, 2, \ldots,$$

Relation (2.4) means roughly that the family either dies out or gets very large. In section 4 it will be shown that (2.4) holds uniformly in r.

<u>Definition</u> 2.4. The random variables $w_n$ are defined by $w_n = z_n/x^n$.

Revised 7-29-48

Clearly $Ew_n = 1$ and $Ew_n^2 = 1 + \frac{\sigma^2}{x^2-x}\left(1-\frac{1}{x^n}\right)$ if $x \neq 1$.

Suppose $n > m$. Then $E(z_n z_m) = \sum_r p_{mr} E(rz_n | z_m = r) = \sum_r p_{mr} r^2 x^{n-m} = x^{n-m} Ez_m^2$. Thus $E(w_n w_m) = Ew_m^2$, whence

$$E(w_n - w_m)^2 = Ew_n^2 - Ew_m^2,\ n > m. \tag{2.5}$$

By virtue of (2.5) we obtain

Theorem 2.1. *If $x > 1$, the random variables $w_n$ converge in mean square, hence in probability, to a random variable $w$.*

For in this case $Ew_n^2 \longrightarrow 1 + \frac{\sigma^2}{x^2-x}$ as $n \longrightarrow \infty$ and (2.5) shows that $E(w_n - w_m)^2 \longrightarrow 0$ as $n$ and $m \longrightarrow \infty$. Theorem 2.1 is then a consequence of [8], p. 38, I.

It is well known that convergence in mean square implies $Ew_n^2 \longrightarrow Ew^2$ and $E(w_n - 1)^2 \longrightarrow E(w-1)^2$ whence $Ew_n \longrightarrow Ew$. Thus we have

$$Ew = 1,\ Ew^2 = 1 + \frac{\sigma^2}{x^2-x}. \tag{2.6}$$

In order to study the behavior of $z_n$ for large $n$ when $x > 1$, we consider the distribution of $w$.

Definition 2.5. $G_n(u) = P(w_n \leq u)$; $\phi_n(s) = E\left(e^{w_n s}\right) = \int_{0-}^{\infty} e^{su} dG_n(u)$.

Definition 2.6. (Applicable when $x > 1$). $G(u) = P(w \leq u)$; $\phi(s) = E(e^{ws}) = \int_{0-}^{\infty} e^{su} dG(u)$. We shall refer to $G(u)$ as *the asymptotic distribution branching from* $f(s)$.

Revised 7-29-48

The moment-generating functions (m.g.f.'s) $\phi_n(s)$ and $\phi(s)$ are defined at least for $Re(s) \leq 0$. Unless specifically stated otherwise we shall consider them only in that domain.

From (2.2) and the fact that $\phi_n(s) = f_n\left[e^{s/x^n}\right]$ it follows that $\phi_{n+1}(sx) = f\left[\phi_n(s)\right]$. Theorem 2.1 implies that if $x > 1$ $G_n(u) \longrightarrow G(u)$ and $\phi_n(s) \longrightarrow \phi(s)$ for $Re(s) \leq 0$. Thus <u>the m.g.f.</u> $\phi(s)$ <u>satisfies the functional equation</u>

$$\phi(sx) = f\left[\phi(s)\right] , \quad Re(s) \leq 0. \tag{2.7}$$

Equation (2.7), which of course is applicable only when $x > 1$, was obtained in a different form by Ulam and Hawkins. It belongs to a type usually known as Koenig's equation, after the nineteenth century mathematician who studied it in connection with functional iteration, and is related to an equation studied by Abel. We shall make some use of the work of Koenig's later. See Hadamard [9] and Koenig [10].

We note that $Ew^k < \infty$ if and only if $Ez_1^k < \infty$. It was already pointed out that $Ew = 1$. As pointed out in [4], as many further moments of $w$ as exist may be found by successive differentiation of (2.7) at $s = 0$.

Finally we note that $G_n(0) = p_{no}$. Hence $\lim G_n(0) = a$. Thus $G(0) = P(w=0) \geq a$. We show later that $G(0) = a$. Clearly $G(u) = 0$ for $u < 0$.

<u>In sections</u> 3 <u>and</u> 4 <u>we always assume</u> $x > 1$.

3. <u>Asymptotic properties of the moment-generating function</u>. We first show that (2.7) uniquely determines the distribution of $w$.

Specifically,

Theorem 3.1. Let $G_1(u)$ and $G_2(u)$ be distributions with equal first moments and finite second moments whose characteristic functions $\phi_1(it)$ and $\phi_2(it)$ satisfy (t is real) $\phi_r(itx) = f[\phi_r(it)]$, $r = 1,2$. Then $G_1(u) = G_2(u)$.

From [13], p. 27, $\phi_1(it) - \phi_2(it) = t^2\beta(t)$, where $\beta(t)$ is bounded as $t \longrightarrow 0$. From (2.7), $\left|\phi_1(itx) - \phi_2(itx)\right| = \left|f[\phi_1(it)] - f[\phi_2(it)]\right| \leq x\left|\phi_1(it) - \phi_2(it)\right|$, since $\left|f'(s)\right| \leq x$ when $|s| \leq 1$. Hence for $t \neq 0$, $\left|\beta\left(\frac{t}{x}\right)\right| \geq x\left|\beta(t)\right|$. Thus $\beta(t)$ cannot be bounded near $t = 0$ unless it is identically zero; hence $\phi_1(it) = \phi_2(it)$.

It is clear that the requirement that $\phi(s)$ have the form $1 + s + O(s^2)$ between two rays from the origin is sufficient for the uniqueness in that domain of solutions of (2.7). On the other hand, continuous solu ons can be constructed at will if the existence of a derivative n 0 is not required.

Before ng further, it is convenient to define three functions $k(s)$, $\Psi(s)$, and $H(u)$ which are closely related to $f(s)$, $\phi(s)$, and $G(u)$ respectively. We repeat that we are considering only the case $x > 1$. See definition 2.3 for a.

Definition 3.1. Let $k(s) = \dfrac{f[s(1-a)+a]-a}{1-a}$. Clearly $k(s)$ is a probability generating function with $k(0) = 0$, $k'(1) = f'(1) = x$, $k''(1) < \infty$. We write $k(s) = \sum_{r=1}^{\infty} q_r s^r$. We also define the iterates $k_n(s)$ by $k_0(s) = s$, $k_{n+1}(s) = k[k_n(s)]$.

Definition 3.2. Let $H(u)$ be the asymptotic distribution branching from $k(s)$. (See Definition 2.6). Let $\Psi(s)$ be the corresponding moment-generating function. We know then that $\Psi(s)$ and $k(s)$ satisfy

$$(3.1) \qquad \Psi(sx) = k\left[\Psi(s)\right].$$

In view of the uniqueness theorem we have, by direct substitution in (3.1), that $\Psi(s)$ must be given by

$$(3.2) \qquad \Psi(s) = \frac{\phi\left[(1-a)s\right] - a}{1-a},$$

and that $H(u)$ must be given by

$$(3.3) \qquad H(u) = \frac{G\left(\frac{u}{1-a}\right)-a}{1-a}, \ u \geq 0; \ H(u) = 0, \ u < 0.$$

We shall see later that $H(0) = 0$; i.e., that $G(0) = a$. Therefore $H(u)$ is the conditional distribution of $(1-a)w$, given that $w \neq 0$. Another way of stating this is as follows:

Theorem 3.2. The random variable $w$ is distributed as the product of two independent random variables $w_o \cdot w'$, where $w_o$ takes the values $0$ and $\frac{1}{1-a}$ with probabilities $a$ and $1-a$ respectively while $w'$ has the asymptotic distribution branching from $k(s)$.

For it is directly verifiable that $\Psi(s)$ is the m.g.f. of $w_o \cdot w'$.

In Theorems 3.3 and 3.4 we consider the behavior of $\Psi(s)$ for large $|s|$. To make for smoother reading we defer the proofs till section 9, where somewhat more general formulations are given. In section 4 the properties of $\Psi(s)$ are interpreted in terms of $G(u)$.

Definition 3.3. Let $\gamma = \log_x\left(\frac{1}{q_1}\right) = \log_x\left[\frac{1}{f'(a)}\right]$. (See definitions 2.3 and 3.1). If $q_1 = 0$ (i.e., $p_o = p_1 = 0$) we take $\gamma = \infty$.

<u>Theorem</u> 3.3. <u>Suppose</u> $\gamma < \infty$. <u>Then</u> <u>if</u> $Re(s) \leq 0$ <u>and</u> $s \neq 0$,

$$\Psi(s) = \frac{M(s)}{|s|^{\gamma}} + M_0(s). \tag{3.4}$$

$M(s)$ <u>is</u> <u>continuous</u> for $s \neq 0$; $M(s)$ <u>and</u> $M_0(s)$ <u>satisfy</u> <u>respectively</u>

$$M(sx) = M(s); \quad M_0(s) = O\left(\frac{1}{|s|^{2\gamma}}\right), \ |s| \longrightarrow \infty. \tag{3.5}$$

<u>Remarks</u>. (See section 9 for proof.) (a). Under the conditions of the theorem $M(s)$ is real and positive when $s$ is real and negative. (b) If $Ez_1^r < \infty$ and the conditions of the theorem hold, the rth derivative of $\Psi(s)$ satisfies

$$\left|\Psi^{(r)}(s)\right| = O\left(\frac{1}{|s|^{r+\gamma}}\right), \ |s| \longrightarrow \infty. \tag{3.6}$$

(c) If $\gamma = \infty$, $\Psi(s)$ and as many derivatives as exist approach 0 exponentially as $|s| \longrightarrow \infty$.

We now consider the behavior of $\Psi(s)$ on the positive real axis, provided it is defined there.

<u>Lemma</u> 3.1. <u>Let</u> $f(s)$ <u>be</u> <u>analytic</u> <u>in</u> <u>the</u> <u>circle</u> $|s| < \alpha, \alpha > 1$. <u>Then</u> $\phi(s)$ <u>and</u> $\Psi(s)$ <u>are</u> <u>analytic</u> <u>in</u> <u>some</u> <u>neighborhood</u> <u>of</u> $s = 0$.

We use a theorem of Poincaré [11] which insures that there is exactly one function $\bar{\phi}(s)$ analytic near $s = 0$ with $\bar{\phi}(0) = \bar{\phi}'(0) = 1$ and satisfying $\bar{\phi}(sx) = f[\bar{\phi}(s)]$. (Although Poincaré's proof is for the case $f(s)$ rational, it applies equally well here.) The circle of convergence of the MacLaurin series for $\bar{\phi}(s)$ has radius $t_\alpha$ where $\bar{\phi}(t_\alpha) = \alpha$. An argument whose details are given in [12], p.21, then

shows that $\phi(s) = \bar{\phi}(s)$ for $|s| < t_\alpha$, and Lemma 3.1 follows. (The argument is necessary to rule out the possibility that the $\phi_n(s)$ converge to $\bar{\phi}(s)$ for $Re(s) \leq 0$ but to some other function for $Re(s) > 0$.) Clearly $\phi(s)$ and $\psi(s)$ are entire if and only if $f(s)$ is entire.

Lemma 3.1 is useful for actual computation of $G(u)$. The (non-negative) coefficients $c_r$ in the series $\phi(s) = 1 + s + c_2 s^2 + \ldots$ can be determined by differentiating (2.7) at $s = 0$. The series can be used to compute values of the characteristic function $\phi(it)$ on some interval $t_0 \leq t \leq t_0 x$, where $t_0$ is a small real number; the values of $\phi(it)$ for the remaining values of $t$ are determined by (2.7). (Note that the real and imaginary parts of $\phi(it)$ are respectively even and odd.) Then the usual inversion formula is used to obtain $G(u)$. A numerical example of this procedure is worked out in section 5.

Definition 3.4. The number $\rho$ is defined by $\rho = \log_x d$ if $f(s)$ is a polynomial of degree $d$, $\rho = \infty$ otherwise.

Theorem 3.4. *Let $f(s)$ (and hence $k(s)$) be a polynomial of degree $d$. Then for $s > 0$*

$$\frac{\log \psi(s)}{s^\rho} = L(s) + L_0(s);$$

*$L(s)$ is continuous and positive; $L(s)$ and $L_0(s)$ satisfy respectively*

$$L(sx) = L(s); \qquad L_0(s) = O\left(\frac{1}{s^\rho}\right), \quad s \to \infty.$$

The proof is in section 9. (Theorem 3.4 may be compared with a more widely applicable but less precise result due to Shah [19]).

<u>Corollary</u>. <u>If</u> $f(s)$ <u>is a polynomial of degree</u> $d$, $\Psi(s)$ <u>is an entire function of order</u> $\rho$ <u>and type</u> $C$ <u>where</u> $C = \operatorname{Max} L(s)$, $1 \leq s \leq x$.

An explicit determination for C has not been found. An approximate numerical determination is not difficult; the function $L(s) = \lim_{n \to \infty} \frac{\log k_n[\Psi(s)]}{s^\rho d^n}$ can be determined numerically for a number of values on some convenient interval $s_0 \leq s \leq s_0 x$, and the maximum value approximated. The importance of C will be indicated in the conjecture following Theorem 4.3. We may also mention that the quantity $[\operatorname{Max} L(s) - \operatorname{Min} L(s)]$, $1 \leq s \leq x$, is of some interest. Some numerical work indicates that in certain cases $L(s)$ is at least approximately constant.

4. <u>Some properties of</u> $G(u)$. Since it will be convenient to work with $H(u)$ rather than $G(u)$, we state the content of Theorems 4.1, 4.2, and 4.3 in terms of $G(u)$: $G(u) = a + \int_0^u g(v)dv$ for $u > 0$. The density $g(u)$ is continuous for $u \neq 0$. If $Ez_1^k < \infty$ then $g^{(r)}(u)$ is continuous for $u \neq 0$ provided $r < \gamma + k - 1$ and is continuous for $u = 0$ provided $r < \gamma - 1$. Near $u = 0$, $G(u)$, provided $\gamma < \infty$, approximates, in a certain mean sense made clear by Theorem 4.2, the function $a + \frac{(1-a)^{\gamma+1}}{\Gamma(1+\gamma)} u^\gamma M[u(1-a)]$, where for convenience we have defined $M(u)$ for positive $u$ by $M(u) = M(-u)$. It is then shown that in a certain sense $g(u)$ goes to zero faster than $\exp(-u^{Q-\varepsilon})$ and slower than $\exp(-u^{Q+\varepsilon})$ where $\varepsilon$ is any positive number, Q being defined in Theorem 4.3. A conjecture is given of a more precise result, applicable when $f(s)$ is a polynomial: in the same sense $g(u)$ goes to zero (more,less) rapidly than $(\exp[-(A^*-\varepsilon)u^Q], \exp[-(A^*+\varepsilon)u^Q])$, where $A^*$ is defined in the conjecture.

Revised 7-29-48

Definition 4.1. Let $H'(u) = h(u)$.

Theorem 4.1. *$H(u)$ is absolutely continuous.* Theorem 3.3 shows that $H(u)$ is continuous; see [13], p. 25. This incidentally shows that $G(0) = a$. If $\gamma > \frac{1}{2}$ the absolute continuity of $H(u)$ follows from the Plancherel theorem. See any text on Fourier transforms. In any case, define the functions $h_m(u) = \frac{1}{2\pi}\int_{-m}^{m} e^{-itu}\psi(it)dt$, $m=1,2,\ldots$. An integration by parts[1] gives for $u \neq 0$

$$(4.1) \qquad h_m(u) = \frac{-1}{2\pi iu}\left[\psi(im)e^{-imu} - \psi(-im)e^{imu}\right] + \frac{1}{2\pi iu}\int_{-m}^{m} e^{-itu}\,\frac{d\psi(it)}{dt}\,dt.$$

If $0 < u_1 \leq u \leq u_2$, (4.1), (3.4), and (3.6) show that the continuous functions $h_m(u)$ converge uniformly in $[u_1, u_2]$ to a continuous function $h(u)$. Moreover

$$(4.2) \qquad H(u_2) - H(u_1) = \lim_{m\to\infty}\int_{-m}^{m}\frac{(e^{-itu_2} - e^{-itu_1})}{-2\pi it}\psi(it)dt$$

$$= \lim_{m\to\infty}\int_{u_1}^{u_2} h_m(u)du = \int_{u_1}^{u_2} h(u)du,$$

the first equality in (4.2) following from [13], p. 28 and the second from the fact that the $h_m(u)$ are uniformly bounded for $u_1 \leq u \leq u_2$. In case $Ez_1^k < \infty$ and $r < \gamma + k - 1$, repeated integration by parts of (4.1) and reference to remark (b), Theorem (3.3), shows that the first

1 I am indebted to J. W. Tukey for this suggestion, which simplifies the original proof.

Revised 7-29-48

r derivatives of $h(u)$ are continuous if $u \neq 0$. The usual integral expression for $h(u)$ in terms of $\Psi(it)$ shows that $\gamma > r+1$ implies $h^{(r)}(u)$ is continuous at 0.

<u>Corollary</u> to the continuity of $H(u)$ : <u>the numbers</u> $p_{nr} = P(z_n = r) \longrightarrow 0$ <u>uniformly in</u> r, $r \geq 1$, <u>as</u> $n \longrightarrow \infty$. We have $p_{nr} = \left[G_n\left(\frac{r}{x^n}\right) - G\left(\frac{r}{x^n}\right)\right] + \left[G\left(\frac{r}{x^n}\right) - G\left(\frac{r}{x^n} - \frac{1}{x^n}\right)\right] + \left[G\left(\frac{r-1}{x^n}\right) - G_n\left(\frac{r-1}{x^n}\right)\right]$. The desired result follows because $G_n(u) \longrightarrow G(u)$ uniformly for $u \geq 0$ and because $G(u)$ must be uniformly continuous for $0 \leq u < \infty$ (right-continuity at 0).

We next consider the behavior of $H(u)$ near $u = 0$, when $\gamma < \infty$. Theorem 3.3 suggests what sort of result may be expected. If the function $M(s)$ of Theorem 3.3 were a constant $M$ it would follow from a Tauberian theorem due to Karamata (see [14], pp. 189-192) that $H(u) \sim \frac{Mu^{\gamma}}{\Gamma(\gamma+1)}$ as $u \longrightarrow 0+$, or $\frac{H(u)}{u^{\gamma+1}} \sim \frac{M}{u\Gamma(\gamma+1)}$. Integrating both sides of this relation from u to ux would give

$$(4.3) \qquad \int_u^{ux} \frac{H(v)dv}{v^{\gamma+1}} \sim \frac{1}{\Gamma(\gamma+1)} \int_1^x \frac{Mdv}{v}.$$

The analogue of (4.3) turns out to be true, as shown by Theorem 4.2, which shows that in a certain mean sense, $H(u)$ behaves like $\frac{u^{\gamma} M(u)}{\Gamma(\gamma+1)}$ as $u \longrightarrow 0+$. (We defined $M(u) = M(-u)$ for $u > 0$.)

<u>Theorem</u> 4.2. $$\lim_{u \longrightarrow 0+} \int_u^{ux} \frac{H(v)dv}{v^{\gamma+1}} = \frac{1}{\Gamma(\gamma+1)} \int_1^x \frac{M(v)dv}{v}$$

The proof, which follows directly along the lines of the proof of Karamata's theorem, is sketched briefly in section 9, for a somewhat

more general situation.

A corollary of Theorem 4.2 is that if $\gamma < 1$, h(u) cannot be bounded as $u \to 0+$; for $h(u) < K$ implies

$$\lim_{u \to 0+} \int_u^{ux} \frac{K \cdot v dv}{v^{\gamma+1}} > \frac{1}{(\gamma+1)} \int_1^x \frac{M(v)}{v} dv > 0, \text{ or } \lim_{u \to 0+} u^{1-\gamma} > 0, \text{ which}$$

implies $\gamma \geq 1$. An example to be given in section 5 shows that if $\gamma = 1$, h(u) is at least in certain cases bounded but discontinuous at 0.

In order to consider the behavior of H(u) as $u \to \infty$ we first prove a theorem which applies to any distribution whose m.g.f. is an entire function.

Theorem 4.3[1] Let F(u) be any c.d.f. whose m.g.f. $\xi(s)$ is entire. Let $\rho$ be the order of $\xi(s)$. Let Q be defined by

$$Q = \text{l.u.b. } q: \int_{-\infty}^{\infty} e^{|u|^q} dF(u) < \infty.$$

Then $\frac{1}{\rho} + \frac{1}{Q} = 1.$

The proof is given in section 9.

Combining Theorems 3.4 and 4.3, we obtain immediately

Theorem 4.4. Let $Q = \text{l.u.b. } q: \int_0^{\infty} e^{u^q} h(u) du < \infty$. Then $Q = \frac{\rho}{\rho - 1}$.

Here $\rho$ is given by definition 3.4. If f(s) is not a polynomial, whether entire or not, the proof of theorem 4.3 will show that Q = 1,

---

1 Before completing the present proof, the writer communicated this result to P. P. Boas, Jr., who sent back a proof along different lines.

Revised 7-29-48

and we interpret theorem 4.4 in that sense. The trivial case $f(s) = s^k$ is excluded, so $\rho > 1$.

Conjecture. Let $\xi(s)$ of Theorem 4.3 be of finite order $\rho$ and type $C$, $0 < C < \infty$. Let $Q = \frac{\rho}{\rho - 1}$ and let $A = \text{l.u.b. } A' : \int_{-\infty}^{\infty} e^{A'|u|^Q} dF(u) < \infty$. Then $(C\rho)^Q \cdot (AQ)^\rho = 1$.

The proof for the case $\rho$ rational follows the same lines as the proof of Theorem 4.3; a general proof has not been found. If the conjecture is true then having determined $\rho$ and $Q$, when $k(s)$ is a polynomial, and having estimated $C$ by the procedure indicated following the corollary to theorem 3.4, we obtain

$$(4.4) \qquad A = \frac{1}{Q}\left(\frac{1}{C\rho}\right)^{\frac{1}{\rho-1}}$$

for the l.u.b. of the numbers $A'$ such that $\int_0^{\infty} e^{A'u^Q} h(u)du < \infty$.

The corresponding number $A^*$ which applies to $g(u)$ is given by

$$(4.5) \qquad A^* = A(1-a)^Q.$$

5. Some special cases. In this section we shall discuss some special cases in which the m.g.f. $\phi(s)$ and the c.d.f. $G(u)$ may be determined explicitly. For these cases and for certain others there is a close relationship between the simple discrete branching process and another type of model to be discussed in section 8. Finally a numerical computation of the distribution $G(u)$ will be given for a particular case where $f(s)$ is a second degree polynomial.

Suppose $f(s)$ has the form

$$f(s) = 1 - \frac{\gamma}{\alpha} + \frac{\gamma}{\alpha}\left(\frac{1}{1+\alpha-\alpha s}\right)$$

with $x > 1$, $\alpha \geq x - 1$, where $f'(1) = x$ and $f''(1) + f'(1) = Ez_1^2 = x(1+2\alpha)$. It is easily verified (as pointed out by Poincaré in [11]), that the solution of the equation $\phi(sx) = f[\phi(s)]$ is given by $\phi(s) = 1 + \frac{(x-1)s}{x-1-\alpha s}$ with $\phi(0) = \phi'(0) = 1$. The number $\underline{a}$ satisfying $a = f(a)$ is given by $a = \frac{\alpha+1-x}{\alpha}$. The functions $\Psi(s)$ and $k(s)$ of section 4 are given by $\Psi(s) = \frac{1}{1-s}$, $k(s) = \frac{s}{x-(x-1)s}$. The number $\gamma$ of Theorem 3.3 is 1. The density function $h(u)$ (definition 4.1) is simply $e^{-u}$, as seen by direct calculation. The number Q of Theorem 4.3 is 1, as it should be, since $f(s)$ is not an entire function. The c.d.f. $H(u)$ is $1-e^{-u}$, and $H(u) \sim u$ near $u = 0$, in agreement with Theorem 4.2. Various aspects of the case $f(s) = \frac{As + B}{Cs + D}$ have been discussed by numerous authors.

Somewhat more generally, we may consider generating functions of the form

(5.1) $$k(s) = s\left[x-(x-1)s^m\right]^{-1/m}, \quad x > 1.$$

The function $k(s)$ is a generating function if and only if m is a non-negative integer. In this case we have $\phi(s) = \Psi(s) = (1-ms)^{-1/m}$ and $g(u) = h(u) = \frac{1}{m^{\frac{1}{m}}\Gamma\left(\frac{1}{m}\right)} u^{\frac{1}{m}-1} e^{-\frac{u}{m}}$. Here $\gamma = \frac{1}{m}$, and we note that unless $m = 1$ the density function $h(u)$ is unbounded near $u = 0$. A physical interpretation for this case will be given in section 8.

As a numerical illustration we consider the case $f(s) = 0.4s + 0.6s^2$. We have $x = Ez_1 = 1.6$ and $\sigma^2 = E(z_1-x)^2 = 0.24$. For the asymptotic distribution, $Ew = 1$, $E(w-1)^2 = \frac{\sigma^2}{x^2-x} = 0.25$.

The number $\gamma = \log_{1.6}\left(\frac{1}{0.4}\right) = 1.9495$ so that $\Psi(s)$ which is identical with $\emptyset(s)$ in this case, is $O\left(\frac{1}{|s|^{1.9495}}\right)$ as $|s|$ goes to $\infty$ with $Re(s) \leq 0$. This implies that the c.d.f. H(u) and likewise G(u), since the two are equal here, behaves like $\left[1/\Gamma(1+\gamma)\right] M(u)$ times $u^{1.9495}$ near u = 0, where the "behavior" is in the sense of Theorem 4.2. Numerical determination of M(u) would not be difficult. The number $\rho$ of Theorem 4.4 is given by $\log_x 2 = 1.4748$. This means that $\Psi(s)$ is an entire function of order 1.4748 and hence that the density function h(u) goes to zero more rapidly than $e^{-u^{Q-\epsilon}}$ and less rapidly that $e^{-u^{Q+\epsilon}}$ for any $\epsilon > 0$, where $Q = \frac{\rho}{\rho - 1} = 3.1061$, and "more rapidly" is used in the sense of Theorem 4.4.

The function $L(s) = \lim_{n \to \infty} \frac{\log \emptyset(sx^n)}{s^{\rho} 2^n}$ was computed for four values of s between s = 1 and s = x = 1.6; in each case the value was 0.744625 so that it appears likely that in this case L(s) is constant. Hence C = Max L(s) = 0.744625 and the quantity A defined by (4.4) is 0.26430. Thus the conjecture following theorem 4.4 indicates that

$$\int_0^{\infty} g(u)\, e^{(0.744625 \pm \epsilon)\, u^{3.1061}}\, du$$ is (divergent, convergent) according as the + or - sign holds.

Through the kindness of Mr. Cecil Hastings of the Douglas Aircraft Company, the c.d.f. G(u) was computed for this case. The coefficients in the power series expansion of $\emptyset(s)$ were obtained from the functional equation (3.1) and G(u) was then obtained by inverting $\emptyset(it)$. The values of G(u) are given in Table I.

# T A B L E I

G(u), the limiting probability that $z_n/x^n \leq u$ for the case $f(s) = 0.4s + 0.6s^2$

| u | G(u) |
|---|---|
| 0.00 | .00000 |
| 0.25 | .04753 |
| 0.50 | .17275 |
| 0.75 | .34550 |
| 1.00 | .53117 |
| 1.25 | .69932 |
| 1.50 | .83042 |
| 1.75 | .91357 |
| 2.00 | .96781 |
| 2.50 | .99751 |
| 3.00 | .99993 |

(Revised) 7/29/48

6. Number of generations to extinction. It was pointed out in section 2 that when $x \leq 1$ the probability is 1 that $z_n = 0$ for some integer n. We assume through-out section 6 that $x < 1$.

Definition 6.1. Let the random variable N be the smallest integer n such that $z_{n+1} = 0$. Define the moment-generating function of N by

$$\Theta(s) = \sum_{n=0}^{\infty} e^{ns} P(N=n).$$

Clearly $P(N=n) = p_{n+1,o} - p_{no}$, so that $\Theta(s) = \sum_{n=0}^{\infty} e^{ns}(p_{n+1,o} - p_{no})$.

Definition 6.2. Let $b_n = 1 - p_{n+1,o}$, with $b_o = 1 - p_o$. The numbers $b_n$ satisfy the recursive relation

$$b_{n+1} = 1 - f[1-b_n]. \tag{6.1}$$

Define the function $\Theta_1(s)$ by

$$\Theta_1(s) = \sum_{n=0}^{\infty} b_n e^{ns}.$$

We see that

$$\Theta(s) = 1 + (e^s - 1)\Theta_1(s), \tag{6.2}$$

so that it suffices to determine the function $\Theta_1(s)$.

The function $\Theta_1(s)$ belongs to a type which has been studied by Fatou [15] and Lattès [16]. If we let $e^s = z$ we see that $\Theta_1(z)$ is a power series whose coefficients are successive iterates of the function $f^*(b) = 1-f(1-b)$; i.e., $b_{n+1} = f^*(b_n) = f^*_{n+1}(b_o)$, where $f^*(0) = 0$, $f^{*\prime}(0) = x < 1$. It was shown by Fatou that a function of this sort is

meromorphic with poles at $s = -n \log x$, $n=1,2,\ldots$. An expansion for $\Theta_1(s)$ in the form

$$\Theta_1(s) = \frac{\mu_1 y_o}{1-xe^s} + \frac{\mu_2 y_o^2}{1-x^2 e^s} + \frac{\mu_3 y_o^3}{1-x^3 e^s} + \ldots,$$

was obtained by Lattès, the expansion converging everywhere except at the poles. The quantities $\mu_r$ and $y_o$ are defined as follows: the function $\mu(s) = \mu_1 s + \mu_2 s^2 + \mu_3 s^3 + \ldots$ is determined by the functional equation $\mu(sx) = f^*[\mu(s)]$ with the condition $\mu'(1) = \mu_1 = 1$. The number $y_o$ is determined by $\mu(y_o) = b_o = 1 - p_o$. Perhaps the easiest way to determine $y_o$ is to use the fact that the inverse function $\mu^{-1}(s)$ satisfies the functional equation $\mu^{-1}[f^*(s)] = x\mu^{-1}(s)$, from which we can determine the power series for $\mu^{-1}(b_o)$.

Since the use of Lattès' expansion requires finding the expansions of $\mu(s)$ and $\mu^{-1}(s)$, we now give another method, giving a different kind of expansion; this method appears particularly adapted to the case here illustrated, where $f(s)$ is of the second degree. Then (6.1) becomes

(6.3) $$b_{n+1} = xb_n - p_2 b_n^2, \quad b_o = 1 - p_o.$$

__Definition__ 6.3. The functions $\Theta_k(s)$, $k=1,2,\ldots$, are given by

(6.4) $$\Theta_k(s) = \sum_{n=0}^{\infty} \left(b_n\right)^k e^{ns}.$$

If we raise both sides of (6.3) to the $k^{th}$ power, multiply both sides by $e^{ns}$, sum on n from 0 to $\infty$, and solve for $\Theta_k(s)$, we obtain

(6.5) $$\Theta_k(s) = \frac{b_o^k e^{-s} + \sum_{r=1}^{k} \binom{k}{r} \left(-p_2\right)^r x^{k-r} \Theta_{k+r}(s)}{e^{-s} - x^k}.$$

(Revised) 7/29/48

(Justification for the rearrangement of series will come out of the subsequent proof.) If we put $k = 1$ in (6.5) we obtain

$$\Theta_1(s) = \frac{b_o e^{-s} - p_2\Theta_2(s)}{e^{-s} - x}. \tag{6.6}$$

Definition 6.4. We define recursively sequences of functions $S_n(s)$ and $R_n(s)$, such that for each $n$, $\Theta_1(s) = S_n(s) + R_n(s)$. Let $S_1(s) = \dfrac{b_o e^{-s}}{e^{-s}-x}$, $R_1(s) = -\dfrac{p_2\Theta_2(s)}{e^{-s}-x}$. Suppose now that $R_n(s)$ is of the form $A_{n1}\Theta_{n+1}(s) + \ldots + A_{nn}\Theta_{2n}(s)$, the $A_{nj}$ being functions of $s$, $p_2$, and $x$, but not explicitly of $b_o$; while $S_n(s)$ is a rational function of $e^{-s}$, $p_2$, and $x$, and a polynomial of degree $n$ in $b_o$. Now put $k = n + 1$ in (6.5) and substitute the expression obtained for $\Theta_{n+1}(s)$ into $R_n(s)$. Collecting terms we now define $R_{n+1}(s)$ as the sum of terms involving $\Theta_{n+2}(s),\ldots,\Theta_{2n+2}(s)$: $R_{n+1}(s) = A_{n+1,1}\Theta_{n+2}(s)+\ldots+A_{n+1,n+1}\Theta_{2n+2}(s)$; then $S_{n+1}(s) = \Theta_1(s) - R_{n+1}(s)$ is a rational function of $e^{-s}$, $p_2$, and $x$, and a polynomial of degree $n + 1$ in $b_o$.

Theorem 6.1. Let $f(s) = p_o+p_1 s+p_2 s^2$, with $x < 1$. Suppose that $x + p_2 b_o < 1$. Then the functions $S_n(s)$ converge to $\Theta_1(s)$ in a neighborhood of $s = 0$.

The restriction $x + p_2 b_o < 1$ may fail to hold. However this is not a serious restriction; we pick a value of $n$ so that $x + p_2 b_n < 1$. Then $\Theta_1(s) = b_o+\ldots+b_{n-1}e^{(n-1)s} + e^{ns}\Theta_1^*(s)$, where the function $\Theta_1^*(s) = \sum_{j=n}^{\infty} b_j e^{(j-n)s}$ is the same type of function as

(Revised) 7/29/48

$\Theta_1(s)$; theorem 6.1 is then applicable to $\Theta_1^*(s)$.

If the conditions of theorem 6.1 are satisfied, we have

$$(6.7) \qquad \Theta_1(s) = b_o e^{-s}\Big[\Pi_1(s,x) - p_2 b_o \Pi_2(s,x) + 2xp_2^2 b_o^2 \Pi_3(s,x) - p_2^3 b_o^3 (e^{-s} + 5x^3)\Pi_4(s,x) + \dots\Big]$$

where $\Pi_k(s,x) = \prod_{r=1}^{k}\left(\frac{1}{e^{-s}-x^r}\right)$. Since $E(N) = \Theta'(0) = \Theta_1(0)$ and $E(N^2) = \Theta''(0) = 2\Theta_1'(0) + \Theta_1(0)$, we have

$$E(N) = b_o\Big[\Pi_1(0,x) - p_2 b_o \Pi_2(0,x) + 2xp_2^2 b_o^2 \Pi_3(0,x) - p_2^3 b_o^3 (1+5x^3)\Pi_4(0,x) + \dots\Big],$$

$$E(N^2) = -E(N) + 2b_o\Big[\Pi_1'(0,x) - p_2 b_o \Pi_2'(0,x) + 2xp_2^2 b_o^2 \Pi_3'(0,x) - (5x^3+1)\, p_2^3 b_o^3 \Pi_4'(0,x) + p_2^3 b_o^3 \Pi_4(0,x) + \dots\Big]$$

where $\Pi_k'(0,x) = \Pi_k(0,x) \sum_{r=1}^{k} \frac{1}{1-x^r}$.

We now prove that if $x + p_2 b_o < 1$, the expansion (6.7) is valid in some neighborhood of $s = 0$. We shall denote the particular values of $x$, $p_2$, and $b_o$ with which we are dealing by $\bar{x}$, $\bar{p}_2$, and $\bar{b}_o$. Now let $x$, $p_2$, and $b_o$ be three complex numbers, arbitrary except for the following restrictions:

$$(6.8) \qquad |x| + |p_2| < 1, \quad |b_o| < 1$$

and define the numbers $b_n$ in terms of $b_o$, $x$, and $p_2$, by means of (6.3), with $\Theta_k(s)$ defined by (6.4).

We first show that (6.7) is valid if (6.8) holds, and then show that the domain of validity also includes the original numbers $\bar{x}$, $\bar{p}_2$, and $\bar{b}_o$, provided $\bar{x} + \bar{p}_2\bar{b}_o < 1$.

If (6.8) is satisfied, we have $|b_n| < A|x|^n$ where A is a positive constant. Now suppose $1 < T < \frac{1}{x}$. Then the series defining $\Theta_k(s)$, $k=1,2,\ldots$, are uniformly and absolutely convergent in the domain $|e^s| \leq T$. Moreover, if $|x| + |p_2| = \Delta < 1$, we have $|b_n| \leq b_o\Delta^n$ whence, if k is an integer large enough so that $T\Delta^k < \frac{1}{2}$,

$$(6.9) \qquad |\Theta_k(s)| \leq 2b_o^k$$

for $|e^s| \leq T$. <u>In what follows, we assume</u> $|e^s| \leq T$. Now write $\Theta_1(s) = S_n(s) + \sum_{j=1}^{n} A_{nj}(p_2,x,s)\Theta_{n+j}(s)$, where n is large enough so that $T\Delta^n < \frac{1}{2}$. Let $A_n(p_2,x,s) = \max_{1\leq j\leq n} |A_{nj}(p_2,x,s)|$. Passing to the next stage we see that $A_{n+1} \leq A_n + \frac{|A_{n1}|}{e^{-s} - x^n}\Delta^{n+1} \leq A_n\left(1 + \frac{\Delta^{n+1}}{e^{-s} - x^n}\right)$.

Hence the numbers $A_n$ are bounded. This fact, together with (6.9), shows that $\lim_{n\to\infty} R_n(s) = 0$.

Now suppose that $x$ and $b_o$ have their original values $\bar{x}$ and $\bar{b}_o$ while $p_2$ is small enough in absolute value so that $\bar{x} + |p_2| < 1$. In this case $\lim_{n \to \infty} S_n(s) = \Theta_1(s)$. We observe that $S_n(s)$ is a polynomial of degree $n - 1$ in $p_2$ and that $S_{n+1}(s)$ is obtained from $S_n(s)$ by adding a single term of degree $n$ in $p_2$. Thus $\Theta_1(s)$ has been expressed as a power series in $p_2$. Now consider $\Theta_1(s)$ as a function of $p_2$, with $b_o = \bar{b}_o$, $x = \bar{x}$. If $\bar{x} + \bar{b}_o |p_2| < 1$, we have $b_n = O[(\bar{x})^n]$. Thus $\Theta_1(s)$ is analytic in $p_2$ for $|p_2| < \frac{1-\bar{x}}{\bar{b}_o}$ and the expansion in (6.7), being a power series in $p_2$, must be valid when $\bar{x} + \bar{p}_2 \bar{b}_o < 1$.

7. Estimation of parameters. Until now we have assumed that the parameters $p_r$ are known numbers. We may wish, however, to estimate them, having observed the numbers $z_1, z_2, \ldots, z_{n+1}$. In order to get simple maximum likelihood estimates for the $p_r$, it appears necessary to introduce certain auxiliary random variables.

Definition 7.1. Let $z_{mk}$ be the number of individuals in the $m^{th}$ generation who have exactly $k$ descendents in the $(m+1)$st generation. Let $Z_n = 1 + z_1 + \ldots + z_n$.

Theorem 7.1. *Maximum likelihood estimates of $p_r$ and $x$, based on observed values of $z_{mk}$ for $m \leq n$, are respectively.*

$$\hat{p}_r = \sum_{m=0}^{n} z_{mr} / Z_n, \quad \hat{x} = \left( Z_{n+1} - 1 \right) / Z_n.$$

(Note that the estimate $\hat{x}$ involves only $z_1, \ldots, z_{n+1}$.)

(Revised) 7/29/48

If $z_m$ is fixed the joint conditional probability function of $z_{mo}, z_{m1}, \ldots,$ is $\left[(z_m)! \prod_{r=0}^{\infty} p_r^{z_{mr}}\right] \Big/ \prod_{r=0}^{\infty} (z_{mr})!$ . Thus the joint probability function of the $z_{mr}$ for $m=0,1,\ldots,n,$ and $r=0,1,2,\ldots,$ is given by the product of two factors, one of which is independent of the $p_r$, the logarithm of the other being $\sum_{r=0}^{\infty}\left(\sum_{m=0}^{n} z_{mr}\right) \log p_r$. The value of this expression is clearly maximized by taking $p_r = \hat{p}_r$ as given above. Since $\sum_r z_{mr} = z_m$ and $\sum_r r z_{mr} = z_{m+1}$, the quantity $\sum r\hat{p}_r$ gives $\hat{x}$ as above.

Although the estimates $\hat{p}_r$ are the same as we would obtain if we were dealing with $Z_n$ trials from a multinomial distribution with probabilities $p_r$, the joint distribution of the quantities $\sum_{m=0}^{n} z_{mr}$, $r=0,1,\ldots,$ is not multinomial. For example, if $Z_n > 1$ the probability of the event $\left\{\sum_{m=0}^{n} z_{mo} = Z_n, \sum_{m=0}^{n} z_{mr} = 0 \text{ for } r \neq 0\right\}$ is 0.

We shall next show that the estimate $\hat{x}$ is, in a certain sense, consistent.

<u>Theorem</u> 7.2. <u>If</u> $x > 1$, <u>the random variables</u> $Z_{n+1}/Z_n$ <u>converge in probability to the random variable</u> $xV^*$ <u>where</u> $V^* = \frac{1}{x}$ if $w = 0$ <u>and</u> $V^* = 1$ if $w \neq 0$.

If $w \neq 0$ then for all $n, z_n \neq 0$ and $1/Z_n \longrightarrow \infty$ as $n \longrightarrow \infty$. Hence in this case $(Z_{n+1}-1)/Z_n$ converges to $x$ if $Z_{n+1}/Z_n$ does. On the other hand, $P(w=0) = a = P(z_n=0)$ for some $n$, so that if $w = 0$, $Z_{n+1}/Z_n = 1$ with probability 1 for $n$ large enough. Thus we

need only show that $Z_{n+1}/Z_n$ converges to x if $x > 1$ and $w \neq 0$.

We need the following:

<u>Lemma 7.1</u>. <u>If</u> $x > 1$, <u>the random variables</u> $Z_n/x^n$ <u>converge in probability to</u> $\frac{wx}{x-1}$.

Since

$$(7.1) \qquad \frac{wx}{x-1} - \frac{Z_n}{x^n} = \frac{w}{x^{n+1}}\left(\frac{x}{x-1}\right) + \sum_{r=0}^{n} \frac{(w-w_r)}{x^{n-r}},$$

it will be sufficient to show that $\lim_{n\to\infty}\left(\frac{x}{x-1}\right)^2 \frac{1}{x^{2n+2}} E(w^2) = 0$ and $\lim_{n\to\infty} E\left(\sum_{r=0}^{n} \frac{(w-w_r)}{x^{n-r}}\right)^2 = 0$. The truth of the first statement is obvious, since $Ew^2$ is finite. It follows from (2.5) that $E(w_r w_s) = Ew_r^2$ if $s > r$, $E(ww_r) = \lim_{n\to\infty} E(w_n w_r) = Ew_r^2$, whence $E(w-w_r)^2 = \frac{\sigma^2}{(x^2-x)x^r}$ and

$E\left[(w-w_r)(w-w_s)\right] = \frac{\sigma^2}{(x^2-x)x^s}$ if $s > r$. Then $E\left(\sum_{r=0}^{n} \frac{(w-w_r)}{x^{n-r}}\right)^2 =$

$\frac{1}{x^{2n}} \frac{\sigma^2}{x^2-x}\left[\sum_{r=0}^{n} x^r + 2\sum_{s=1}^{n}\sum_{r=0}^{s-1} x^r\right]$, and this quantity clearly approaches 0 as $n \to \infty$, proving Lemma 7.1.

Define the random variables $w^*$ and $V_n$ as

$$w^* = w \quad \text{when} \quad w \neq 0$$
$$w^* = 1 \quad \text{when} \quad w = 0$$
$$V_n = \frac{Z_n}{x^n} \quad \text{when} \quad z_n \neq 0$$
$$V_n = \frac{x}{x-1} \quad \text{when} \quad z_n = 0$$

(Revised) 7/29/48

It is clear that the $V_n$ converge in probability to $w^* \frac{x}{x-1}$, and we note that the c.d.f. of $w^*$ is continuous at $w^* = 0$. Hence,

$$\lim_{n \to \infty} P\left(\left|\frac{V_{n+1}}{V_n} - 1\right| > \mathcal{E} > 0\right) = \lim_{n \to \infty} P\left(V_{n+1} - V_n \pm V_n \mathcal{E} \lesseqgtr 0\right) =$$

$$P\left(\frac{\pm w^* x \mathcal{E}}{x-1} \lesseqgtr 0\right) = 0.$$ It follows, under the conditional hypothesis $w \neq 0$, that the variates $\frac{Z_{n+1}}{Z_n}$ converge in probability to x, since

$\frac{Z_{n+1}}{Z_n} = x \frac{V_{n+1}}{V_n}$ when $z_{n+1} \neq 0$.

8. Continuous Models. As mentioned in section 1 there are situations where it is more important to consider the number of individuals existing at a given time than the number in a given generation. Let a set of probabilities $p_r$ be given. The question arises whether we can interpret these as probabilities that an individual will have a given number of descendents at the end of some fixed period of time. We might then suppose that each individual in existence at that time has the same probabilities of having a given number of descendents at the end of the next (equal) length of time, these probabilities being independent of the age of the individual. A model of this sort might be considered in certain fission processes, if the probability of fission is independent of age. It should be noted that the "descendents" of an individual may include the individual. For example, if a bacterium splits in two we may either regard it as having produced two descendents and dying, or as having produced one descendent and itself surviving.

If an interpretation of this sort is to be satisfactory, interpolation in time must be possible. In other words there should exist a family of functions $f_n(s)$ defined for all positive n such that $f_{n_1}\left[f_{n_2}(s)\right] = f_{n_1+n_2}(s)$; such that for each positive n, $f_n(s)$ is a probability generating function, $f_n(s) = \sum_{r=0}^{\infty} p_r(n)s^r$; and such that for $n = 0,1,2,\ldots$ the functions $f_n(s)$ coincide with the iterates $s, f(s), f\left[f(s)\right], \ldots$. We may then interpret $f_n(s)$ as the generating function at time n. It is readily seen that in general such a family of functions will not exist. For example, if such a family exists we must have $f(s) = n^{th}$ iterate of $f_{1/n}(s)$ for arbitrarily large integral n, so that f(s) cannot be a polynomial of degree $\geq 2$.

The functional equation $\phi(sx) = f\left[\phi(s)\right]$ shows that $f(s) = \phi\left[x\phi^{-1}(s)\right]$, whence $f_n(s) = \phi\left[x^n\phi^{-1}(s)\right]$ for integral n. The expression $\phi\left[x^n\phi^{-1}(s)\right]$ then can be taken as the definition of $f_n(s)$ for all positive n. See Hadamard, [9]. The problem of determining whether the functions so defined are a family of generating functions will be discussed in a subsequent paper. We remark, however, that if f(s) has the form $\frac{s}{x-(x-1)s}$ considered in section 5 then the iterates $f_n(s)$ have the form $\frac{s}{x^n-(x^n-1)s}$; they are clearly generating functions for all positive n, satisfying the required relation $f_{n_1}(f_{n_2}) = f_{n_1+n_2}$. Now suppose g(s) is some function such that the function $f(s) = g^{-1}\left[\frac{g(s)}{x-(x-1)g(s)}\right]$ is a generating function for all $x > 1$, with $g(1) = 1$. As pointed out by

(Revised) 7/29/48

Ulam and Hawkins, the iterates of functions $f(s)$ of this form are convenient to work with, the $n^{th}$ iterate being simply $g^{-1}\left[\frac{g(s)}{x^n-(x^n-1)g(s)}\right]$.

In addition, the requirement that $f(s)$ be a generating function for all $x > 1$ shows that the functions $f_n(s)$ are generating functions for all $n > 0$. The simplest function $g(s)$ which satisfies our requirements is $g(s) = s^m$, where $m$ is any positive integer. In this case $f(s)$ has the form considered in (5.1) and $f_n(s) = s\left[x^n-(x^n-1)s^m\right]^{-\frac{1}{m}}$. As $n \longrightarrow 0$, we have $f_n(s) = (1 - \frac{n}{m}\log x)s + \frac{n\log x}{m}s^{m+1} + O(n^2)$. We may interpret this as follows. A particle in existence at a given time may, in a short time interval $\Delta t$, either split into $m+1$ particles, with probability $\frac{\Delta t \log x}{m}$; or it may remain unaltered, with probability $1 - \frac{\Delta t \log x}{m}$. If it splits, each particle produced has the same chances for splitting as its parent, etc. Thus, from the results of section 5, it follows that if we begin with a single particle at time $t = 0$, the asymptotic probability density function for $z_t/x^t$, where $z_t$ is the number of particles at time $t$, is given by

$$\left(m^{-\frac{1}{m}}u^{\frac{1}{m}-1}e^{-\frac{u}{m}}\right)\Big/\Gamma\left(\frac{1}{m}\right).$$

It is, of course, customary to begin with the elementary probabilities for a certain number of births in a short time $\Delta t$ and determine the functions $f_n(s)$ from these by means of differential equations. See, for example, Arley, [17]. The results of the present paper can be applied in some cases to the continuous problem even when an explicit determination of the $f_n(s)$ is difficult. A discussion will be given in a later paper.

9. <u>Some proofs</u>. We give in this section proofs for (A) theorem 3.3, (B) theorem 3.4, (C) theorem 4.2, and (D) theorem 4.3; in certain cases we shall indicate slightly more general results.

(A) We make use of a result of Koenigs, in the form applicable here.

<u>Koenigs' theorem</u>: <u>If</u> $|s| \leq \lambda < 1$ <u>and</u> $q_1 \neq 0$, <u>then</u> $k_n(s) = q_1^n B(s)\left[1+O(q_1^n)\right]$ <u>where</u> $B(s)$ <u>is analytic for</u> $|s| \leq \lambda$ <u>and satisfies the functional equation</u> $B\left[k(s)\right] = q_1 B(s)$.

Here, $O(q_1^n)$ means bounded by $Aq_1^n$, where A is independent of s. We remark that $B(s) \not\equiv 0$. The proof of Koenigs' theorem follows readily if we write $k_n(s) = q_1^{n-1}k(s)\prod_{j=1}^{n-1}\left\{1 + \frac{\xi\left[k_j(s)\right]}{q_1}\right\}$, where $\xi(s) = \frac{k(s)}{s} - q_1$.

Now let $t_1$ be a positive number such that $\left|\Psi(s)\right| < 1$ when $0 < |s| \leq t_1$ and $Re(s) \leq 0$. (For the rest of this proof we assume $Re(s) \leq 0$.) Such a number exists; on the imaginary axis we have $\Psi(it) = 1 + it - \frac{1}{2}E\left[(w')^2\right]t^2 + o(t^2)$ where $E\left[(w')^2\right] > 1$, w' having the distribution branching from k(s), showing that $\left|\Psi(it)\right| < 1$ if $t \neq 0$ and sufficiently small; while if $Re(s) < 0$ we refer to the expression $\Psi(s) = \int_0^\infty e^{su}dH(u)$. Let $\lambda = \text{Max}\left|\Psi(s)\right|$ for $t_1/x \leq |s| \leq t_1$.
If $|s| > t_1$ let N(s) be the smallest integer such that $|s|/x^{N(s)} \leq t_1$.
Then $\Psi(s) = k_{N(s)}\left[\Psi\left(s/x^{N(s)}\right)\right] = q_1^{N(s)}B\left[\Psi\left(s/x^{N(s)}\right)\right]\left[1 + O(q_1^{N(s)})\right] = B\left[\Psi(s)\right]\left[1 + O(q_1^{N(s)})\right]$. Now $B\left[\Psi(sx)\right] = q_1 B\left[\Psi(s)\right]$.
Let $M(s) = |s|^\gamma B\left[\Psi(s)\right]$. Then $M(sx) = M(s)$. Also $\log_x\left|s/t_1\right| \leq N(s) < 1 + \log_x\left|s/t_1\right|$, and theorem 3.3 follows. Clearly $M(s)/|s|^\gamma$ is

continuous for $t_1 x \leq |s| \leq t_1$ and hence, by functional continuation, wherever $Re(s) \leq 0$, $s \neq 0$.

Concerning the remarks following Theorem 3.3 we have the following:

(a) If $Ez_1^r < \infty$, r-fold differentiation of $\Psi(sx^n) = k_n[\Psi(s)]$ gives, for $|s| \geq t_1 > 0$,

$$\Psi^{(r)}(s) = \frac{1}{x^{nr}} \sum_{j=1}^{r} Q_{rj} k_n^{(j)} \left[\Psi\left(\frac{s}{x^n}\right)\right], \tag{9.1}$$

where $Q_{rj}$ is a polynomial in $\Psi^{(1)}\left(\frac{s}{x^n}\right), \ldots, \Psi^{(r)}\left(\frac{s}{x^n}\right)$. Now $|k_n(s)| = O(q_1^n)$ when $|s| \leq \lambda$; because of analyticity, the same must be true of $|k_n^{(j)}(s)|$. Put $n = N(s)$ in (9.1), $N(s)$ being the integer defined above. Since $k_N^{(j)}[\Psi(s/x^N)] = O(q_1^N) = O\left(1/|s|^{\gamma}\right)$, remark (a) follows.

(b) $B(s)$ is clearly $\geq 0$ when $s \geq 0$; hence $M(s) \geq 0$ when $s < 0$. Since $B(0) = 0$, $B(s) \neq 0$ for sufficiently small $s \neq 0$; since $\Psi(s) \longrightarrow 0$ as $|s| \longrightarrow \infty$, $M(s) \neq 0$ for $|s|$ sufficiently large; since $M(sx) = M(s)$, remark (b) follows.

(c) If $\gamma = \infty$, i.e., $q_1 = 0$, then $k_n(s)$ goes to zero with great rapidity as $n \longrightarrow \infty$, if $|s| < 1$. The general line of argument is clear.

(B) Let $k(s)$ be a polynomial of degree $d > 1$ with real coefficients, $k(s) = q_0 + \ldots + q_d s^d$, with a non-negative double point, $k(\alpha) = \alpha \geq 0$, and such that $k(s) > s$ when $s > \alpha$. Let $\Psi(s)$ be any solution of the functional equation $\Psi(ms) = k[\Psi(s)]$ which is continuous for $s > 0$

(Revised) 7/29/48

<u>and satisfies</u> $\Psi(s) > \alpha$ <u>for</u> $s > 0$; <u>here</u> $m$ <u>is any number</u> $> 1$. <u>Then theorem</u> 3.4 <u>holds, with</u> $x$ <u>replaced by</u> $m$.

It is not difficult to show that if $\alpha < s_1 \leq s \leq s_2$, $\lim_{j\to\infty} k_j(s) = \infty$ uniformly in $s$. Hence $\Psi(s) \to \infty$ as $s \to \infty$. Write $R(s) = \log(1 + \frac{1}{q_d}\sum_{j=1}^{d} q_{d-j}s^{-j})$. Then $d^{-n}\log\Psi(sm^n) = d^{-n}\log k_n\left[\Psi(s)\right] =$

$(1-d^{-n})\log q_d/(d-1) + \log\Psi(s) + \sum_{j=1}^{n} d^{-j}R\left(k_{j-1}\left[\Psi(s)\right]\right)$, $s$ being taken large enough so that $R\left(k_{j-1}\left[\Psi(s)\right]\right)$ is continuous. Thus, since the functions $R(k_{j-1}\left[\Psi(s)\right])$ are bounded, the functions $d^{-n}\log\Psi(sm^n)$ converge uniformly, for $s$ sufficiently large, to a continuous function $L^*(s)$ satisfying $L^*(ms) = dL^*(s)$. Let $L(s) = t^{-\rho}L^*(s)$, where $\rho = \log_m d$.

Theorem 3.4 now follows by an argument similar to that used to conclude theorem 3.3. (Note that $\sum_{n+1}^{\infty} d^{-j}R\left(k_{j-1}\left[\Psi(s)\right]\right) = O(d^{-n})$).

(C) In order to avoid negative signs we work with the Laplace transform instead of the m.g.f.

<u>Let</u> $H(u)$ <u>be nondecreasing on</u> $(0,\infty)$ <u>with</u> $H(0) = 0$; <u>let</u> $\Psi(s) = \int_0^\infty e^{-su}dH(u)$ <u>be finite for</u> $s > 0$. <u>Suppose</u> $\Psi(s) = \frac{M(s)}{s^\gamma} + o(\frac{1}{s^\gamma})$ <u>as</u> $s \to \infty$ where $0 < \gamma < \infty$, $M(s)$ <u>is continuous and satisfies</u> $M(sx) = M(s)$ <u>for</u> $s > 0$, $x$ <u>being some number</u> $> 1$. <u>Then</u> $\lim_{u\to 0+}\int_u^{ux}\frac{H(v)}{v^\gamma}dv =$

$$\frac{1}{\Gamma(\gamma+1)}\int_1^x \frac{M(v)}{v}\,dv.$$

Following the lines of the proof of Karamata's theorem, we see that for any $y > 0$, $\int_y^{xy} s^{\gamma-1}\psi(s)ds = D + o(1)$ as $s \longrightarrow \infty$ where $D = \int_1^x \frac{M(s)}{s}\,ds$; i.e., $\int_y^{xy} s^{\gamma-1}ds \int_0^\infty e^{-su}dH(u) = D + o(1)$, or replacing s by $(n+1)s$, $\int_y^{xy} s^{\gamma-1}ds\int_0^\infty e^{-su}e^{-nsu}dH(u) = D/(n+1)^\gamma + o(1) = \frac{D}{\Gamma(\gamma)}\int_0^\infty e^{-s}e^{-ns}s^{\gamma-1}ds + o(1)$. It follows as in [14], pp. 189-192, that if $F(u)$ is any function of bounded variation in (0,1) we have

$$\text{(9.2)}\qquad \lim_{y\to\infty}\int_y^{xy} s^{\gamma-1}ds\int_0^\infty e^{-su}F(e^{-su})dH(u) = \frac{D}{\Gamma(\gamma)}\int_0^\infty e^{-s}F(s^{-s})s^{\gamma-1}ds.$$

Let $F(e^{-s}) = e^s$ if $0 \leq s \leq 1$ and 0 otherwise. Then the theorem follows from (9.2).

(D) Theorem 4.3 is true if $F(u)$ is any bounded monotone increasing function. For simplicity we assume that $F(1) = 0$; it is readily seen that this causes no loss in generality. The proof is given for the case $1 < \rho < \infty$; it will be clear that $\rho = 1$ implies $Q = \infty$, while if $\rho = \infty$ (or if $\xi(s)$ is not entire) $Q = 1$.

Suppose m and n are positive integers such that $m/n < \rho/(\rho-1)$. Then

$$\text{(9.3)}\qquad \int_1^\infty \exp(u^{\frac{m}{n}})\,dF(u) = \sum_{r=0}^\infty \frac{1}{r!}\int_1^\infty u^{\frac{mr}{n}}dF(u) \leq n\sum_{r=0}^\infty \frac{[(r+1)m]!}{(rn)!}\,c_{(r+1)m}$$

where $c_k = \frac{\xi^{(k)}(0)}{k!}$; interchange of integration and summation are justified by the positiveness of all terms involved. Suppose $0 < \varepsilon < \frac{n}{m} - (1-\frac{1}{\rho})$; for k sufficiently large the inequality $c_k < k^{-k(\frac{1}{\rho}-\varepsilon)}$ is satisfied; see [18], p. 253. Hence, using Stirling's formula, we see that the last series in (9.3) is dominated by a series whose $r^{th}$ term, for r sufficiently large, is controlled by the factor $r^{rm(1-\frac{1}{\rho}+\varepsilon-\frac{n}{m})}$. Since $1-\frac{1}{\rho}+\varepsilon-\frac{n}{m}$ is negative, the series, and hence the integral, converges. We have thus proved $\frac{1}{Q}+\frac{1}{\rho}\leq 1$.

Conversely, suppose $\frac{m}{n} > \frac{\rho}{\rho-1}$. Let $\xi(s) = \sum_{k=0}^{m-1} \xi_k(s)$, where $\xi_k(s) = \sum_{r=0}^{\infty} c_{k+rn} s^{k+rn}$, $k=0,1,\ldots,m-1$. At least one of the functions $\xi_k(s)$ must be of order $\rho$. We suppose that $\xi_0(s)$ is; if not the argument would need only slight modifications. We have

$$(9.4) \qquad \int_1^{\infty} \exp(u^{\frac{m}{n}})\,dF(u) \geq n \sum_{r=0}^{\infty} \frac{(rm)!\,c_{rm}}{[(r+1)n]!} .$$

Suppose $0 < \varepsilon < 1 - \frac{1}{\rho} - \frac{n}{m}$. From [18], p.253, the inequality $c_{rm} > (rm)^{-rm(\frac{1}{\rho}+\varepsilon)}$ must hold for infinitely many values of r. As in the first half of the proof this shows that the series and the integral in (9.4) diverge. Thus $\frac{1}{\rho}+\frac{1}{Q}\geq 1$ and the proof is complete.

If $\rho$ is rational, the conjecture following theorem 4.3 can be proved in a similar manner making use of a relation between the class of an entire function and the coefficients of its series expansion; see [14], p. 95.

## REFERENCES


[1] R. A. Fisher, _The Genetical Theory of Natural Selection_, Oxford, the Clarenden Press, 1930.

[2] Alfred J. Lotka, _Theorie Analytique des Associations Biologiques_, 2, Paris, Herman, 1930.

[3] J. F. Steffensen, "Deux problemes du calcul des probabilites," _Ann. de l'Inst. Henri Poincaré_, Vol. 3 (1933), p. 331.

[4] D. Hawkins and S. Ulam, _Theory of Multiplicative Processes_, I, Los Alamos Declassified Document 265, 1944.

[5] A. Kolmogoroff, "Zur losung einer biologischen aufgabe", _Communications for Math. and Mechanics at Tchebycheff University_, Tomsk, Vol. 2, part 1 (1938), p. 1; see _Fortschritte der Math._, Vol. 64 (1938), part 2, p. 1223.

[6] A. Kolmogoroff and N. A. Dmitriev, "Branching stochastic processes", _C. R. (Doklady) Acad. Sci. URSS (N.S.)_, Vol. 56 (1947) pp. 5-8. See _Math. Rev._ Vol. 9, No. 1 (1948), p. 46.

[7] A. M. Yaglom, "Certain limit theorems of the theory of branching random processes", _Doklady Akad. Nauk SSSR (N.S.)_ Vol. 56 (1947), pp. 795-798. See _Math. Rev._ Vol. 9, No. 3 (1948), p. 149.

[8] A. Kolmogoroff, _Grundbegriffe der Wahrscheinlichkeitsrechnung_, Chelsea, 1946.

[9] J. Hadamard, "Two works on iteration and related questions", _Bull. Am. Math. Soc._, Vol. 50 (1944), p. 67.

[10] G. Koenigs, "Nouvelles recherches sur les integrales de certaines equations fonctionnelles", _Ann. Sci. de l'Ecole Normale Sup. de Paris_, Vol. 1, series 3, suppl., (1884) p. 3.

[11] H. Poincaré, "Sur une classe nouvelle de transcendantes uniformes", _Journal de Math. Pures et Appliquees_, Vol. 6, 4th series (1890), p. 313.

[12] T. E. Harris, _Some Theorems on the Bernoullian Multiplicative Process_, Dissertation, Princeton, 1947.

[13] H. Cramér, _Random Variables and Probability Distributions_, Cambridge Tract 36, 1937.

[14] D. V. Widder, _The Laplace Transform_, Princeton University Press, 1941.

[15] P. Fatou, "Sur une classe remarkable de series de Taylor", _Ann. Sci. de l'Ecole Normale Sup. de Paris_, (3), Vol. 27 (1910), pp.43-53.

REFERENCES (Contd.)

[16] S. Lattès, "Sur les suites recurrentes non lineaires et sur les fonctions generatrices de ces suites", Ann. de la Fac. des Sciences de l'Univ. de Toulouse (3), Vol. 3 (1911), pp. 96-105.

[17] Niels Arley, On the Theory of Stochastic Processes and Their Applications to the Theory of Cosmic Radiation, G.E.C. Gads Forlag, Copenhagen, 1943.

[18] E. C. Titchmarsh, The Theory of Functions, 2d Ed., Oxford, 1939.

[19] S. M. Shah, "On real continuous solutions of algebraic difference equations", Bull. Am. Math. Soc., Vol. 53 (1947), pp. 548-558.